

**ARRANGEMENT OF DATA CELLS AND NEURAL NETWORK SYSTEM
UTILIZING SUCH AN ARRANGEMENT**

Patent Number: ☐ US5075889
Publication date: 1991-12-24
Inventor(s): DURANTON MARC A G (FR); JOUSSELIN CHRISTIAN P M (FR)
Applicant(s): PHILIPS CORP (US)
Requested Patent: ☐ EP0369551, B1
Application Number: US19890439717 19891120
Priority Number(s): FR19880014999 19881118
IPC Classification: G11C13/00
EC Classification: G11C8/12
Equivalents: DE68923763D, DE68923763T, ☐ FR2639461, ☐ JP2173860, ☐ US5151971

Abstract

An arrangement of data cells which stores at least one matrix of data words which are arranged in rows and columns, the matrix being distributed in the arrangement in order to deliver/receive, via a single bus, permuted data words which correspond either to a row or to a column of the matrix. Each data cell is connected to the single bus via series-connected switches which are associated with a respective addressing mode, the switches which address a same word of a same mode being directly controlled by a same selection signal. Circulation members enable the original order of the data on the bus to be restored. An arrangement of this kind is used in a layered neural network system for executing the error backpropagation algorithm. Application: Calculator, microprocessors, processor, neural network system. Reference: FIG. 4.

Data supplied from the esp@cenet database - I2

This Page Blank (uspto)

(12)

DEMANDE DE BREVET EUROPEEN

(21) Numéro de dépôt: 89202899.4

(51) Int. Cl.⁵: G11C 8/00, G11C 5/02

(22) Date de dépôt: 16.11.89

(30) Priorité: 18.11.88 FR 8814999

(43) Date de publication de la demande:
23.05.90 Bulletin 90/21

(84) Etats contractants désignés:
CH DE FR GB IT LI SE

(71) Demandeur: LABORATOIRES
D'ELECTRONIQUE PHILIPS
3, Avenue Descartes
F-94450 Limeil-Brévannes(FR)
(84) FR

Demandeur: N.V. Philips'
Gloeilampfabrieken
Groenewoudseweg 1
NL-5621 BA Eindhoven(NL)

(84) CH DE GB IT LI SE

(72) Inventeur: Jousseiln, Christian
Société Civile S.P.I.D. 209, rue de l'Université
F-75007 Paris(FR)

Inventeur: Duranton, Marc
Société Civile S.P.I.D. 209, rue de l'Université
F-75007 Paris(FR)

(74) Mandataire: Landousy, Christian et al
Société Civile S.P.I.D. 209, Rue de
l'Université
F-75007 Paris(FR)

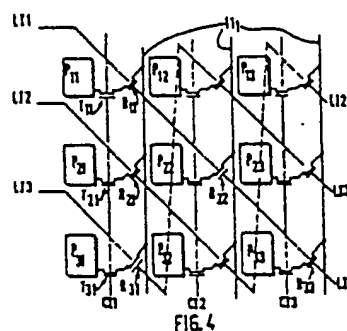
(54) Arrangement de cellules de données et structure de réseaux de neurones utilisant un tel arrangement

(57) Arrangement de cellules de données qui stocke au moins une matrice de mots de données ordonnés selon des lignes et des colonnes, la matrice étant distribuée dans l'arrangement pour délivrer/recevoir sur un bus unique des mots de données permutés correspondant soit à une ligne soit à une colonne de la matrice. Chaque cellule de données est réunie au bus unique à travers des commutateurs mis en série, propres à chaque mode d'adressage, les commutateurs adressant un même mot d'un même mode étant directement contrôlés par un même signal de sélection.

Des organes de circulation permettent de rétablir le rangement original des données sur le bus. Un tel arrangement est utilisé dans une structure de neurones en couches pour mettre en oeuvre l'algorithme

de la rétropropagation des erreurs.

Application : calculateur, microprocesseur, processeur de traitement, structure de réseaux de neurones.



ARRANGEMENT DE CELLULES DE DONNEES ET STRUCTURE DE RESEAUX DE NEURONES UTILISANT UN TEL ARRANGEMENT

L'invention concerne un arrangement de cellules de données qui stocke au moins une matrice de mots de données ordonnés selon des lignes et des colonnes, la matrice étant distribuée dans l'arrangement pour délivrer/recevoir sur un bus unique, à l'aide de moyens de sélection de mode, des mots de données permutés correspondant soit à une ligne soit à une colonne de la matrice.

Elle concerne également une structure de réseaux de neurones utilisant un tel arrangement.

On connaît le document WO 84/00629 qui traite d'un système de mémoire à accès multidimensionnel. Ce document considère le cas de l'écriture et de la lecture d'un tableau de données à deux dimensions stocké dans une mémoire. Pour permettre l'adressage de mots de données soit selon les lignes soit selon les colonnes du tableau, les données sont rangées dans la mémoire d'une manière particulière, afin que celles-ci puissent être lues/entrées à l'aide d'un bus unique. Pour cela les données sont chargées ligne par ligne, la première ligne chargée correspondant à la première ligne du tableau, la deuxième ligne chargée correspondant à la deuxième ligne du tableau mais après avoir subi une permutation circulaire des données de une case mémoire. Chaque ligne chargée est ainsi décalée de 1 case mémoire par rapport à la précédente et subit en conséquence la permutation circulaire correspondante. Les données qui étaient sur une colonne du tableau se trouvent ainsi disposées selon une diagonale de la mémoire. Les données qui étaient sur une ligne du tableau se trouvent toujours disposées sur la ligne de même rang de la mémoire mais avec une permutation circulaire déterminée. Pour adresser une ligne ou une colonne du tableau qui sont distribuées sur plusieurs colonnes de la mémoire, il est alors nécessaire de déterminer à chaque fois l'adresse vraie en mémoire à l'aide d'un modificateur d'adresses. Les données qui apparaissent avec une certaine permutation sont rétablies dans l'ordre initial à l'aide d'un organe qui opère la rotation des données. Chaque modificateur d'adresses affecté à chaque colonne effectue un calcul d'adresses qui dépend à la fois de la ligne et de la colonne du tableau et de la colonne de la mémoire. De tels calculs nécessitent un matériel important qui présente les inconvénients de retarder le calcul des adresses vraies et d'être trop encombrant pour pouvoir permettre une réalisation intégrée compacte. Le problème posé est donc de réaliser un arrangement du genre qui vient d'être décrit et qui soit intégrable d'une manière compacte et qui ait un fonctionnement rapide. La solution consiste en un arrange-

ment pour lequel chaque cellule de données est réunie au bus unique à travers des commutateurs mis en série, propres à chaque mode d'adressage, les commutateurs adressant un même mot d'un même mode étant directement contrôlés par un même signal de sélection.

Ainsi avantageusement une ligne unique portant un unique signal de sélection de mode va pouvoir adresser la totalité d'un mot donné en activant à la fois tous les commutateurs concernés. Chaque commutateur peut être formé d'un transistor de sélection. Ceci accroît la compacité et la rapidité de l'arrangement en particulier lorsque celui-ci est réalisé sous forme de circuit intégré.

La matrice de mots de données peut être distribuée dans l'arrangement selon une distribution de type diagonal de sorte qu'une ligne (une colonne) de la matrice étant distribuée en diagonale dans l'arrangement, les autres lignes (autres colonnes) de la matrice soient distribuées parallèlement à cette diagonale.

Les bits constituant les mots de données apparaissent sur le bus dans un ordre qui est différent de celui qu'ils avaient dans les mots de données de la matrice. Pour rétablir cet ordre il faut alors opérer un rangement qui est effectué par un registre à décalage à circulation commandé par un organe de circulation qui décale les données selon le rang adressé de la matrice. Ce rangement des données sur le bus peut également être effectué par un additionneur qui détermine une adresse interne du circuit tampon en ajoutant une valeur prédéterminée propre au rang adressé dans l'arrangement à une adresse externe du circuit tampon.

Lorsqu'on désire distribuer plusieurs matrices de mots de données, il est possible de le faire dans un arrangement bidimensionnel ou dans un arrangement multidimensionnel. Dans ce dernier cas chaque dimension dispose de son mode de sélection. On peut ainsi avoir trois modes de sélection X, Y, Z dans un arrangement tridimensionnel.

L'intérêt pour une réalisation compacte et rapide se manifeste en particulier dans les systèmes qui nécessitent de stocker un grand nombre de tableaux qui peuvent être de grandes dimensions. C'est par exemple le cas des systèmes de réseaux de neurones. Il est nécessaire alors de disposer d'une intégration la plus dense possible avec des vitesses de traitement élevées.

Les réseaux de neurones sont formés d'automates élémentaires connectés entre eux par des synapses auxquelles sont associés des coefficients synaptiques. Un réseau de neurones met en œu-

vre des phases de résolution au cours desquelles il détermine des états V_j de neurones de sortie j à partir des états V_i de neurones d'entrée i . Cette phase de résolution s'opère selon :

$$V_j = \sum_i C_{ij} V_i$$

Mais pour s'adapter à une tâche déterminée, le réseau de neurones doit également opérer des phases d'apprentissage. Un tel réseau de neurones est par exemple décrit dans le document : "A chipset for high speed simulation of neural network systems" S.C.J. GARTH IEEE Conference on Neural Networks SAN DIEGO III-443 (1987). Les phases d'apprentissage ont pour objet de modifier les coefficients synaptiques afin que le réseau s'adapte à un problème déterminé.

Pour la mise en oeuvre de l'algorithme d'apprentissage selon la rétropropagation de l'erreur, le réseau de neurones présente une structure en couches. Au cours d'une phase de type "résolution", des états de neurones de sortie sont déterminés à l'aide des coefficients synaptiques stockés dans le réseau. Puis un ordinateur hôte compare les états de neurones de sortie déterminés avec les états de neurones attendus. Les erreurs observées sont introduites dans la dernière couche de la structure puis rétropropagées avec un sens de propagation inversé par rapport au sens du mode de résolution. Pour mettre en oeuvre successivement une phase de résolution puis une phase d'apprentissage on utilise les coefficients synaptiques selon une matrice C_{ij} dans la phase de résolution et selon une matrice C_{ji} , transposée de la matrice C_{ij} , dans la phase d'apprentissage comme cela est indiqué dans le document de S.C.J. GARTH.

Ceci nécessite d'effectuer à chaque changement de phases des opérations répétées de chargement des moyens de stockage des coefficients synaptiques par la matrice soit C_{ij} soit la transposée C_{ji} .

Ces opérations de chargement nécessitent beaucoup de temps pour être effectuées.

La quantité de calculs à effectuer étant très importante, il est possible d'accroître la rapidité en mettant en oeuvre un traitement parallèle. Ceci nécessite de pouvoir lire/écrire la matrice C_{ij} par blocs de lignes ou de colonnes selon la phase des opérations.

Pour cela on peut chercher à doubler le nombre de moyens de stockage des coefficients synaptiques ou le nombre de bus mais cela nuit à la compacité de la structure et nécessite des moyens de sélection supplémentaires qui retardent néanmoins le fonctionnement.

C'est pourquoi selon l'invention on utilise l'arrangement de cellules de données décrit précédemment. Dans ce cas l'invention concerne une structure de réseaux de neurones en couches qui comprend :

- des moyens de résolution pour déterminer pour chaque couche des états de neurones de sortie i à partir des états de neurones d'entrée j reliés entre eux par des synapses,

- des moyens de stockage des coefficients synaptiques C_{ij} propres à ces synapses,

- des moyens d'apprentissage et de réactualisation des coefficients synaptiques C_{ij} ,

- des moyens de stockage des états de neurones, caractérisée en ce que les moyens de stockage des coefficients synaptiques sont constitués d'au moins un arrangement déjà décrit, qui stocke au moins une matrice carrée de coefficients synaptiques C_{ij} qui sont distribués dans au moins une mémoire, celle-ci possédant un double adressage et un bus unique de sorte qu'un mot de coefficients synaptiques relatif soit à une ligne soit à une colonne de la matrice carrée peut être adressé et placé sur le bus unique, afin que la structure de réseaux de neurones opère des phases d'apprentissage soit avec la matrice de coefficients synaptiques soit avec la matrice transposée pour mettre en oeuvre l'algorithme de rétropropagation des erreurs.

Selon une disposition préférentielle les coefficients synaptiques sont distribués dans une mémoire bidimensionnelle selon une distribution de type diagonal de sorte qu'une ligne (une colonne) de la matrice étant distribuée selon une diagonale de la mémoire, les autres lignes (autres colonnes) de la matrice soient distribuées parallèlement à cette diagonale.

Ainsi avantageusement les mots de coefficients synaptiques apparaissent sur le bus dans un ordre prédéterminé et restent néanmoins faciles à exploiter selon l'ordre initial à l'aide d'un registre à décalage à circulation. Pour cela les moyens de stockage des états de neurones comprennent un registre à décalage à circulation qui permet de faire circuler les états de neurones pour apparier chaque état de neurone à son coefficient synaptique.

Ainsi au lieu de remettre les coefficients synaptiques dans leur ordre initial on fait subir aux vecteurs d'états de neurones les permutations circulaires appropriées.

Ces permutations circulaires sont contrôlées par un ordinateur d'adresses qui opère la circulation dans le registre à décalage en générant les commandes correspondant au rang du mot de coefficient synaptique adressé.

Comme cela peut être déduit de ce qui a déjà été indiqué, l'invention concerne en général un système de manipulation des données comprenant une pluralité de cellules muni d'une sélection pour communiquer des données avec un environnement, le système de manipulation de données étant muni de moyens de sélection pour sélection-

ner un groupe de cellules pour réaliser des couplages parallèles entre les cellules sélectionnées et un unique bus de données à travers des chemins de données respectifs. Chaque cellule particulière appartient au moins respectivement à un premier groupe et respectivement à un second groupe, qui ont seulement la cellule particulière en commun et qui peut être sélectionnée respectivement selon un premier et un second mode. Pour cela les moyens de sélection transmettent au moins un premier et un second signal de sélection qui contrôlent le chemin de données entre chaque cellule et le bus de données.

Ainsi l'accès aux cellules est permis selon des configurations prédéterminées contrôlés par les deux signaux de sélection. Dans le cas où la pluralité de cellules a été organisée selon une matrice principale, les modes de sélection peuvent permettre par exemple une communication de données selon des rangées en parallèle concernant les rangées de l'organisation matricielle et une communication de données en parallèle concernant les diagonales de par exemple une sousmatrice carrée de la matrice principale.

Ceci est avantageux dans les dispositifs concernant le traitement de données réalisé à la fois sur une matrice de données et sur sa transposée. Comme cela a été décrit auparavant, un dispositif de cette sorte se présente dans les réseaux de neurones adaptatifs en couches qui sont soumis à un apprentissage selon l'algorithme connu de rétropropagation. En général l'avantage apparaît dans l'initialisation d'une matrice de données unique qui est stockée, la communication de données en parallèle s'opérant à grande vitesse selon un mécanisme de sélection simple.

Comme cela a déjà été indiqué, chaque chemin de données peut comprendre un arrangement série de commutateurs chacun étant contrôlé par un signal de sélection respectif. On peut aisément prévoir une caractéristique de l'invention dans laquelle ce chemin de données comprend un arrangement série de canaux de conduction de transistors qui constituent les commutateurs. Un équivalent pratique peut comprendre un transistor contrôlé par un signal de détection supplémentaire représentatif d'une fonction logique desdits premier et second signaux de sélection et générés au moyen d'une porte logique propre à chaque cellule, réduisant ainsi les seuils et les pertes de saturation dans les transistors des chemins de données. Dans le cas où une cellule stocke un mot à plusieurs bits cette architecture peut économiser une certaine surface d'intégration d'un circuit intégré.

Cette structure de réseaux de neurones peut constituer un circuit intégré compact opérant un traitement parallèle rapide.

L'invention sera mieux comprise à l'aide des

figures suivantes données à titre d'exemple non limitatif qui représentent :

figure 1 : une matrice représentant un exemple de rangement de données selon une disposition classique des données.

figure 2 : une matrice représentant un exemple de disposition des données selon l'art antérieur.

figure 3A : un exemple de réalisation des moyens de rangement des données de sortie.

figure 3B : un autre exemple de réalisation des moyens de rangement des données de sortie.

figure 4 : une structure d'adressage de la matrice de points mémoire à deux dimensions.

figure 5A : une partie d'un circuit de réseaux de neurones selon l'invention.

figures 5B, 5C : deux dispositions des données au cours de phases de résolution et de phases de rétropropagation du gradient.

figures 6A, 6B : une structure d'adressage à trois dimensions.

Pour la clarté de l'exposé la description est faite avec un arrangement à deux dimensions de points mémoire. Sa généralisation à trois dimensions est présentée sur les figures 6A et 6B. Des dimensions d'ordre supérieur s'en déduisent sans difficulté.

La figure 1 représente une matrice 10 dans laquelle les données sont stockées selon un adressage ligne ou colonne habituel. Pour permettre une explication claire les nombres mis dans chaque case représentent les données stockées. Dans ce cas les lignes/colonnes de la matrice coïncident avec les lignes/colonnes de l'arrangement. Ainsi la ligne L1 contient les données 00 jusqu'à 70, et la colonne C1 contient les données 00 jusqu'à 07. La matrice peut être adressée soit en ligne soit en colonne. Ainsi, si la colonne C3 est adressée les données de la colonne apparaissent sur le bus 11, et peuvent être stockées dans le registre tampon 12₁. La donnée 20 est dans la case supérieure du registre et la donnée 27 est dans la case inférieure du registre. De même si la ligne L3 est adressée, les données sélectionnées apparaissent sur le bus 11₂ et peuvent être stockées dans un registre tampon 12₂. La donnée 02 est dans la case gauche du registre et la donnée 72 est dans la case droite du registre. Ainsi les données des lignes ou des colonnes apparaissent sur des bus différents ce qui nécessite des circuits tampons différents. Si avec une telle disposition on cherche à utiliser un seul circuit tampon ceci ne peut se faire qu'au prix d'interconnexions supplémentaires, certaines assez longues, et avec ajout d'éléments de sélection nombreux.

Selon le document de l'art antérieur, la matrice de mots de données va être disposée dans l'arrangement de cellules de données selon une disposition qui est représentée sur la figure 2. Contraire-

mment au cas précédent il y a lieu de distinguer une ligne/rangée de la matrice et une ligne/rangée de l'arrangement. Sur cette disposition on observe que les données qui étaient sur la première ligne de la figure 1 sont placées sur la diagonale de la matrice carrée de la figure 2. Chaque colonne est alors rangée en mettant la valeur de la première ligne de la figure 1 dans la case de la figure 2 placée sur la diagonale, puis les autres valeurs, par exemple en se déplaçant vers le bas, et lorsque le bas de la matrice est atteint on continue le chargement de la colonne par la case de la première ligne. Selon cet arrangement lorsque la colonne C3 est adressée, les données de la colonne apparaissent sur le bus de sortie puis éventuellement dans le circuit tampon 12₁ selon le rangement représenté par I. On observe que les données de sortie sont les mêmes que celles de la figure 1 mais avec un rangement différent. Ainsi la donnée 20 apparaît non pas dans la case supérieure mais dans la troisième case du circuit tampon 12₁. Ainsi chaque donnée est décalée d'un nombre de positions égal au numéro d'ordre de la colonne adressée.

D'une manière analogue lorsque la ligne L3 est adressée les données sélectionnées apparaissent sur le même bus et dans le même circuit tampon 12₁ que précédemment selon le rangement représenté par II. Le circuit tampon 12₁ est représenté deux fois sur la figure 2 pour la commodité de l'exposé. On observe également que les données sont les mêmes que lorsque la ligne L3 est adressée selon la disposition de la figure 1 mais là encore avec un rangement différent. Chaque donnée est ainsi décalée d'un nombre de positions égal au numéro d'ordre de la ligne adressée.

L'exemple présenté est une matrice carrée mais l'invention s'applique à des matrices rectangulaires MxN à condition que le bus unique et si nécessaire le circuit tampon unique soient placés sur le côté de la matrice qui délivre le maximum d'éléments binaires. Les éléments mémoires qui possèdent le même adressage en diagonale s'obtiennent par translation de la matrice carrée soit MxM soit NxN inscrite dans la matrice rectangulaire MxN.

La figure 4 représente la structure d'adressage de l'arrangement de cellules de données selon l'invention qui exploite la disposition de la matrice de données représentée sur la figure 2. Dans cet exemple simple correspondant à un arrangement de 3x3 les points P11, P21, P31 d'une même colonne sont sélectionnés par un même signal de sélection C11 qui opère sur des transistors T11, T21, T31. De même les signaux C12, C13 sélectionnent les autres colonnes. Les points mémoires P11, P22, P33 situés sur une même diagonale de l'arrangement (même colonne de la matrice) sont sélectionnés par le signal L11 qui opère sur les

transistors R11, R22, R33. Le signal L12 opère sur les points P21, P32 et P13. Le signal L13 opère sur les points P31, P12 et P23. Pour opérer la sélection d'un point il faut rendre actif par exemple soit tous les signaux L1 et un des signaux C1 ou soit l'inverse. Les données sont délivrées en sortie sur le bus 11₁. Inversement pour écrire il suffit de présenter les données sur le bus 11₁, d'effectuer l'adressage et de mettre les points mémoires en état d'écriture.

L'exemple qui vient d'être traité concerne un arrangement à deux dimensions. Les figures 6A et 6B représentent le cas d'un arrangement de cellules à trois dimensions. Dans ce cas selon l'invention (fig. 6A) il est possible d'effectuer l'adressage selon par exemple une direction D correspondant à une dimension de la matrice et une direction diagonale par exemple les directions E ou F. La matrice tridimensionnelle étant repérée par trois indices, la direction D correspond à des points ayant un des indices constant. Lorsque cet adressage dans la direction D concerne uniquement la partie de matrice de la face avant du cube représenté, (un seul indice constant), alors les groupements de cellules sélectionnés correspondent à une colonne de cellules de cette face avant. Lorsque cet adressage dans la direction D concerne la totalité de l'espace tridimensionnel (deux indices constants) alors les groupements de cellules sélectionnés correspondent à une tranche basée sur ladite colonne et qui s'étend dans la direction perpendiculaire à ladite face avant. La matrice tridimensionnelle est représentée pour la commodité de l'explication sous la forme d'un cube mais peut présenter une forme quelconque. En particulier l'arrangement multidimensionnel peut être topologiquement disposé selon une surface (deux dimensions).

Le second adressage va pouvoir être effectué selon les directions diagonales E ou F pour lesquelles soit deux soit trois indices restent identiques entre eux. Cela concerne donc les diagonales elles-mêmes et les lignes de cellules qui leur sont parallèles.

Chacune des dimensions de la matrice possède des moyens d'adressage représentés sur la figure 6B qui se déduisent de la structure de la figure 4. Un seul élément mémoire P314 est représenté. Les transistors T314 (commande C1₁), S314 (commande D1₄), R314 (commande L1₃) permettent d'effectuer l'adressage selon les trois dimensions de la matrice. Un point mémoire par exemple P314 est sélectionné en rendant actives par exemple toutes les commandes C1 et D1 et seulement la commande L1₃. En effet pour activer une cellule quelconque il faut rendre active la totalité de l'adressage selon deux dimensions de la matrice et selon l'autre dimension seulement la commande qui concerne ladite cellule.

Les figures 3A et 3B représentent une matrice MxN 10 munie d'un bus unique 11₁ réuni à un circuit tampon 12₁ qui peut stocker des tranches de M bits. La figure 3A représente une première variante de réalisation des moyens de rangement des données de sortie. Pour cela, à l'aide d'une commande externe 19, un générateur 20 délivre des adresses qui déterminent l'adressage ligne 21₁ et l'adressage colonne 21₂ respectivement à l'aide de décodeurs 25₁ et 25₂. Cette même commande 19 arrive à un organe de circulation 22 qui selon le rang de l'adresse délivrée, soit ligne soit colonne, délivre une commande de circulation 23 au circuit tampon 12₁ constitué d'un registre à décalage bouclé sur lui-même. Ainsi les données qui ont été lues à l'aide de l'adressage opéré puis stockées dans le circuit tampon 12₁ sont mises dans l'ordre souhaité par cette opération avant d'être délivrées sur le bus 13.

La figure 3A permet de sortir les données en parallèle. Il peut être intéressant de pouvoir les sortir en série selon le rangement désiré. Pour cela selon la figure 3B la mise en ordre des données de sortie se fait directement à l'aide d'un décodage d'adresses. Pour cela la commande 19 de l'adressage de la mémoire 10 arrive sur un additionneur 31 qui reçoit également une adresse externe 35 destinée à lire le circuit tampon 12₁ qui dans ce cas n'est plus un registre à décalage bouclé mais une ligne de points mémoire réunis à une sortie série unique 36. A l'adresse 35 est additionnée une valeur contenue dans la commande 19 qui dépend du rang de la ligne ou de la colonne adressée dans la mémoire 10. L'additionneur est réuni à un décodeur qui délivre l'adresse interne du point mémoire désiré dans le circuit tampon 12₁. L'adresse externe 35 peut provenir d'un dispositif externe ou d'un compteur qui détermine les adresses pour la lecture du circuit tampon 12₁.

La figure 5A représente une partie d'un circuit de réseau de neurones traitant un exemple simplifié dans le cas d'un arrangement à deux dimensions. Pour un neurone 1, l'arrangement 10 délivre sur son bus les quatre coefficients synaptiques C₁₁, C₁₂, C₁₃, C₁₄ qui arrivent sur des multiplieurs 50₁, 50₂, 50₃, 50₄ qui reçoivent respectivement les états de neurones V₁, V₂, V₃, V₄ qui sont stockés dans le registre à décalage 12₁ qui constitue dans ce cas la mémoire d'états de neurones. Ces multiplieurs sont réunis à un arbre d'additionneurs 51 qui délivre un résultat S tel que :

$$S = C_{11}.V_1 + C_{12}.V_2 + C_{13}.V_3 + C_{14}.V_4.$$

Pour un neurone 2 l'arrangement va délivrer les coefficients synaptiques avec un rangement différent tel que représenté sur la figure 5B. Avant d'effectuer le calcul pour la phase de résolution, le registre à décalage 12₁ fait circuler les états de neurones pour appairier correctement chaque état

de neurone à son coefficient synaptique. Le coefficient synaptique C₂₁ est apparié avec l'état V₁, le coefficient synaptique C₂₂ est apparié avec l'état V₂ et ainsi de suite.

La figure 5C concerne la phase d'apprentissage. En effet dans ce cas les coefficients synaptiques sont lus selon la matrice transposée : C₁₁, C₂₁, C₃₁, C₄₁. Dans ce cas le registre 12₁ contient non pas les états de neurones mais les erreurs D₁, D₂, D₃, D₄ afin de mettre en oeuvre l'algorithme de rétropropagation des erreurs. Les mêmes mécanismes de circulation sont mis en oeuvre pour appairer les données. Ainsi le registre habituel qui stocke les états de neurones est transformé en un registre à décalage à circulation pour pouvoir s'adapter aux mots de coefficients synaptiques permutés délivrés par l'arrangement.

Mais le mécanisme de circulation peut ne pas être effectué par le registre à décalage, les coefficients étant mémorisés dans des registres temporaires chargés à partir de multiplexeurs à N voies (pour N coefficients synaptiques) qui rétablissent l'ordre correct des coefficients.

Dans les cas où l'on ne fait pas un accès aléatoire des neurones mais un accès séquentiel et dans l'ordre: neurone 1, puis 2, puis 3 etc..., après chaque évaluation de neurones, on opère un décalage dans le registre à décalage contenant soit les états de neurones soit les erreurs, et ainsi ces termes sont toujours en correspondance avec les coefficients synaptiques qui les pondèrent. Au bout de N calculs, on retombe dans la position initiale de départ.

Dans un circuit où les états sont codés sur m bits, et où les multiplieurs font aussi des multiplications par des termes de m bits, le registre à décalage est constitué de m registres à décalages identiques, chacun portant un des poids de la représentation des états.

Dans un circuit où les états sont codés sur m bits, et où les multiplieurs font des multiplications sur 1 bit (avec des portes ET ou OU-exclusif par exemple), il faut m cycles pour pouvoir calculer la somme pondérée pour un neurone. Dans ce cas le registre à décalage comporte N * m positions et il effectue un décalage après chaque calcul de bit.

Revendications

1. Arrangement de cellules de données qui stocke au moins une matrice de mots de données ordonnés selon des lignes et des colonnes, la matrice étant distribuée dans l'arrangement pour délivrer/recevoir sur un bus unique, à l'aide de moyens de sélection de mode, des mots de données permutés correspondant soit à une ligne soit à une colonne de la matrice, caractérisé en ce que

chaque cellule de données est réunie au bus unique à travers des commutateurs mis en série, propres à chaque mode d'adressage, les commutateurs adressant un même mot d'un même mode étant directement contrôlés par un même signal de sélection.

2. Arrangement selon la revendication 1 caractérisé en ce que chaque commutateur comprend un transistor de sélection.

3. Arrangement selon une des revendications 1 ou 2 caractérisé en ce que la matrice de mots de données est distribuée dans l'arrangement selon une distribution de type diagonal de sorte qu'une ligne (une colonne) de la matrice étant distribuée en diagonale dans l'arrangement, les autres lignes (autres colonnes) de la matrice soient distribuées parallèlement à cette diagonale.

4. Arrangement selon une des revendications 1 à 3 caractérisé en ce que le rangement des données d'un mot sur le bus est effectué par un registre à décalage à circulation commandé par un organe de circulation qui décale les données selon le rang adressé de la matrice.

5. Arrangement selon une des revendications 1 à 4 caractérisé en ce que le rangement des données d'un mot sur le bus est effectué par un additionneur qui détermine une adresse interne du circuit tampon en ajoutant une valeur prédéterminée, propre au rang adressé dans l'arrangement, à une adresse externe du circuit tampon.

6. Arrangement selon une des revendications 1 à 5 caractérisé en ce que plusieurs matrices de mots de données sont distribuées dans l'arrangement selon plusieurs dimensions correspondant à plusieurs modes de sélection.

7. Arrangement selon une des revendications 1 à 6 réalisé sous forme de circuit intégré.

8. Structure de réseaux de neurones en couches qui comprend :

- des moyens de résolution pour déterminer pour chaque couche des états de neurones de sortie i à partir des états de neurones d'entrée j reliés entre eux par des synapses,
- des moyens de stockage des coefficients synaptiques C_{ij} propres à ces synapses,
- des moyens d'apprentissage et de réactualisation des coefficients synaptiques C_{ij} ,
- des moyens de stockage des états de neurones, caractérisée en ce que les moyens de stockage des coefficients synaptiques sont constitués d'au moins un arrangement selon une des revendications 1 à 7, qui stocke au moins une matrice carrée de coefficients synaptiques C_{ij} qui sont distribués dans au moins une mémoire, celle-ci possédant un double adressage et un bus unique de sorte qu'un mot de coefficients synaptiques relatif soit à une ligne soit à une colonne de la matrice carrée peut être adressé et placé sur le bus unique, afin que la

structure de réseaux de neurones opère des phases d'apprentissage soit avec la matrice de coefficients synaptiques soit avec la matrice transposée pour mettre en oeuvre l'algorithme de rétropropagation des erreurs.

9. Structure de réseaux de neurones selon la revendication 8 caractérisée en ce que les coefficients synaptiques sont distribués dans une mémoire bidimensionnelle selon une distribution de type diagonal de sorte qu'une ligne (une colonne) de la matrice étant distribuée en diagonale dans la mémoire, les autres lignes (autres colonnes) de la matrice soient distribuées parallèlement à cette diagonale.

10. Structure de réseaux de neurones selon une des revendications 7 ou 9 caractérisée en ce que les moyens de stockage des états de neurones comprennent un registre à décalage à circulation qui permet de faire circuler les états de neurones pour apparier chaque état de neurone à son coefficient synaptique.

11. Structure de réseaux de neurones selon la revendication 10 caractérisée en ce qu'elle comprend un calculateur d'adresses qui opère la circulation dans le registre à décalage en générant les commandes correspondant au rang du mot de coefficients synaptiques adressé.

12. Structure de réseaux de neurones selon une des revendications 8 à 11 réalisée sous forme de circuit intégré.

13. Système de manipulation de données comprenant une pluralité de cellules muni d'une sélection pour communiquer des données avec un environnement, le système de manipulation de données étant muni de moyens de sélection pour sélectionner un groupe de cellules pour réaliser des couplages parallèles entre les cellules sélectionnées et un unique bus de données, caractérisé en ce que chaque cellule particulière appartient au moins respectivement à un premier groupe qui est sélectionné selon un premier mode de sélection et respectivement à un second groupe qui est sélectionné selon un second mode de sélection, lesdits groupes ayant seulement la cellule particulière en commun, les moyens de sélection transmettant pour cela au moins un premier et un second signal de sélection contrôlant en conséquence chaque chemin de données entre chaque cellule et le bus de données.

14. Système de manipulation de données selon la revendication 13 dans lequel la pluralité de cellules est organisée selon une matrice principale d'au moins deux dimensions caractérisé en ce que le premier groupe comprend des cellules qui sont arrangées selon une première sous-matrice unidimensionnelle comprise dans la matrice principale, et le second groupe comprend des cellules qui forment une diagonale d'au moins une unique se-

conde sous-matrice d'au moins deux dimensions ayant un nombre égal de cellules pour toutes ses dimensions et étant comprise dans la matrice principale.

15. Système de manipulation de données selon une des revendications 13 ou 14 caractérisé en ce que chaque chemin de données comprend un arrangement série de commutateurs, chacun étant contrôlé par un signal de sélection respectif. 5

16. Système de manipulation de données selon une des revendications 13 ou 14 caractérisé en ce que le chemin de données pour chaque cellule comprend un commutateur qui est contrôlé par un signal de sélection supplémentaire représentant une fonction logique desdits premier et second signaux de sélection. 10 15

17. Système de manipulation de données selon la revendication 14 caractérisé en ce que la pluralité de cellules constitue une mémoire pour stocker une matrice de données, des moyens de réarrangement étant couplés audit bus de données en rapport avec une configuration prédéterminée réarrangeant les données transmises par lesdits couplages parallèles selon le premier mode de sélection pour permettre une communication de données en lignes ou en colonnes par rapport à la matrice de données et selon le second mode de sélection pour permettre une communication de données en lignes ou en colonnes par rapport à la matrice de données transposée, ladite configuration dépendant du mode sélectionné. 20 25 30

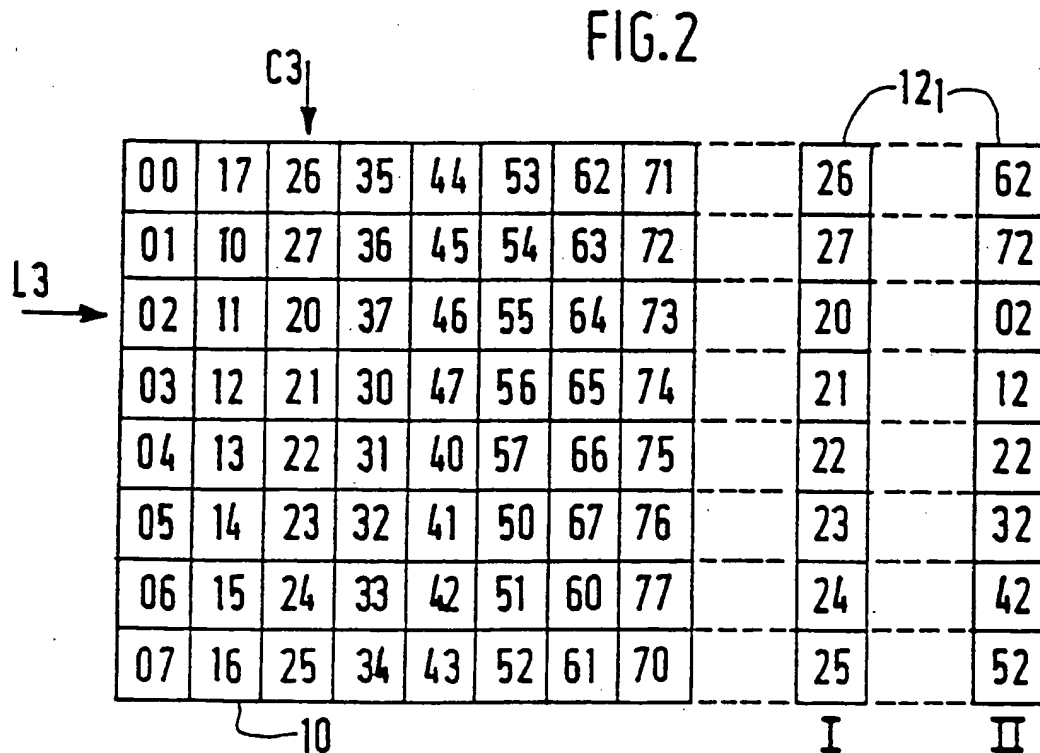
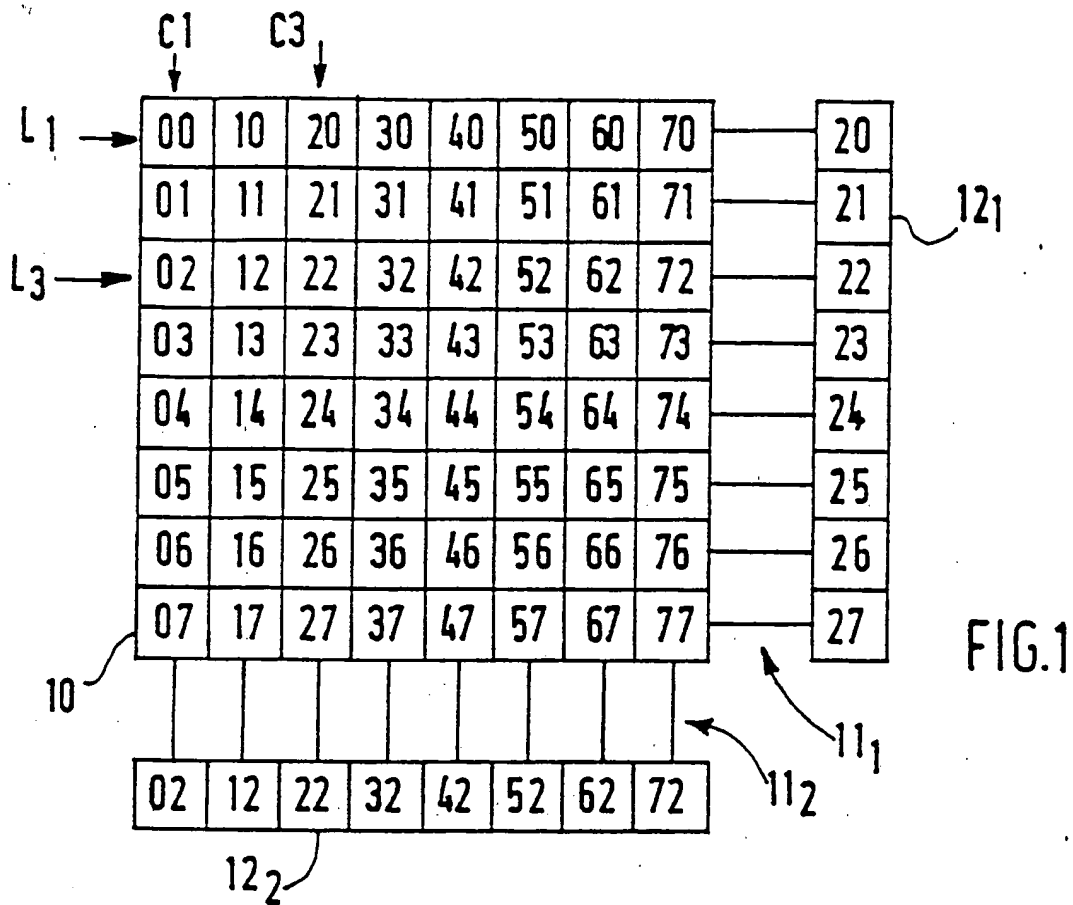
18. Réseau de neurones comprenant un système de manipulation de données selon la revendication 17, la mémoire stockant une matrice de coefficients synaptiques. 35

40

45

50

55



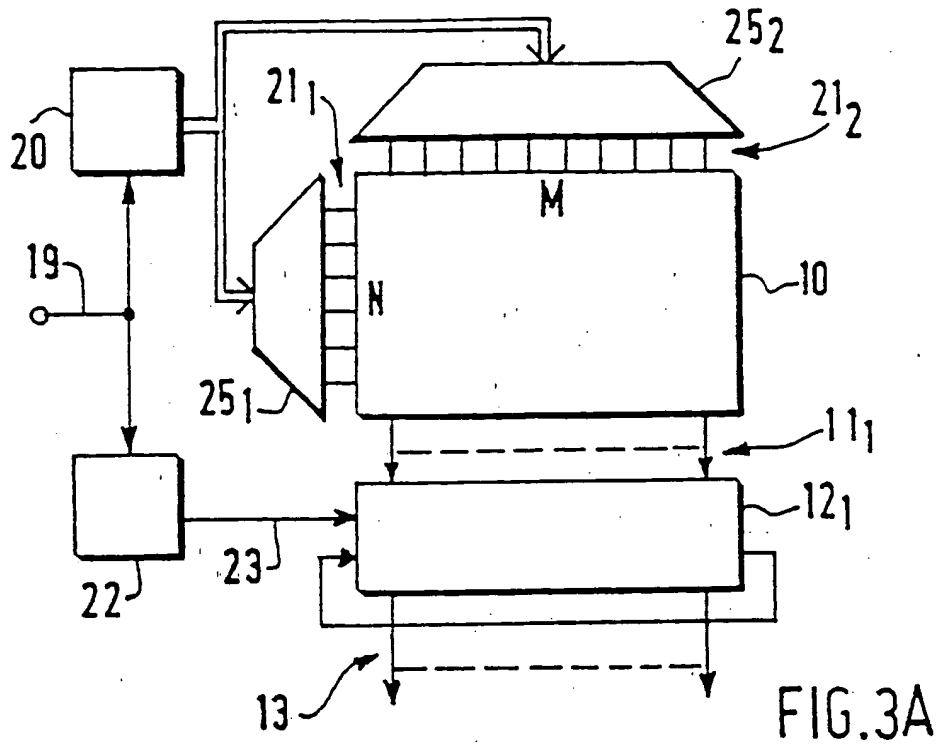


FIG. 3A

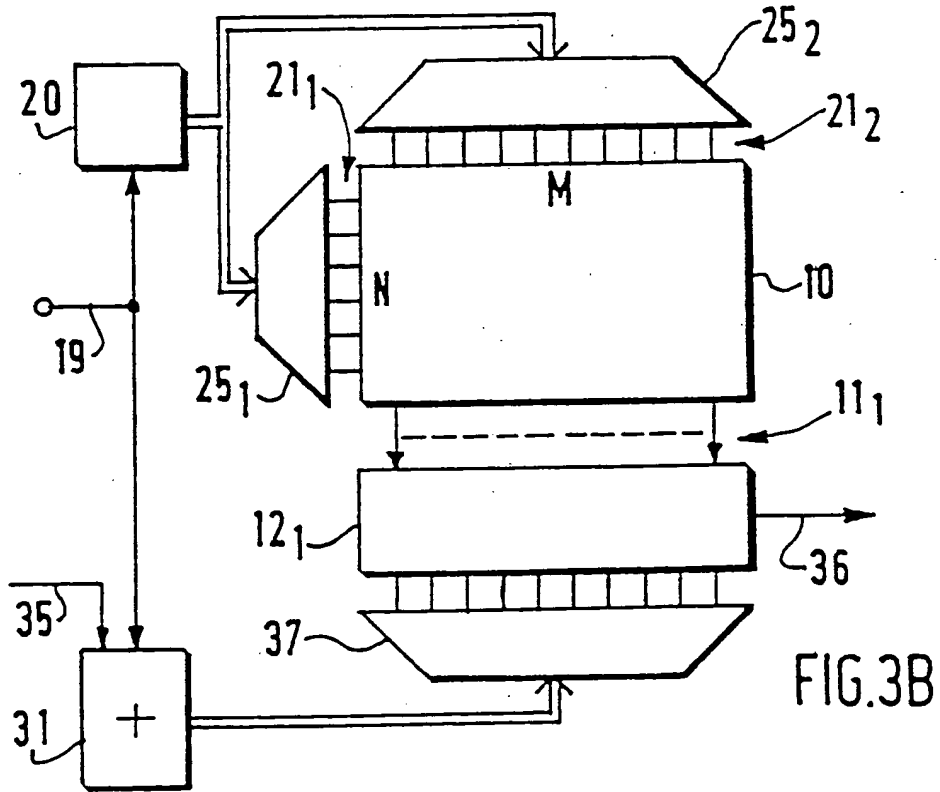


FIG. 3B

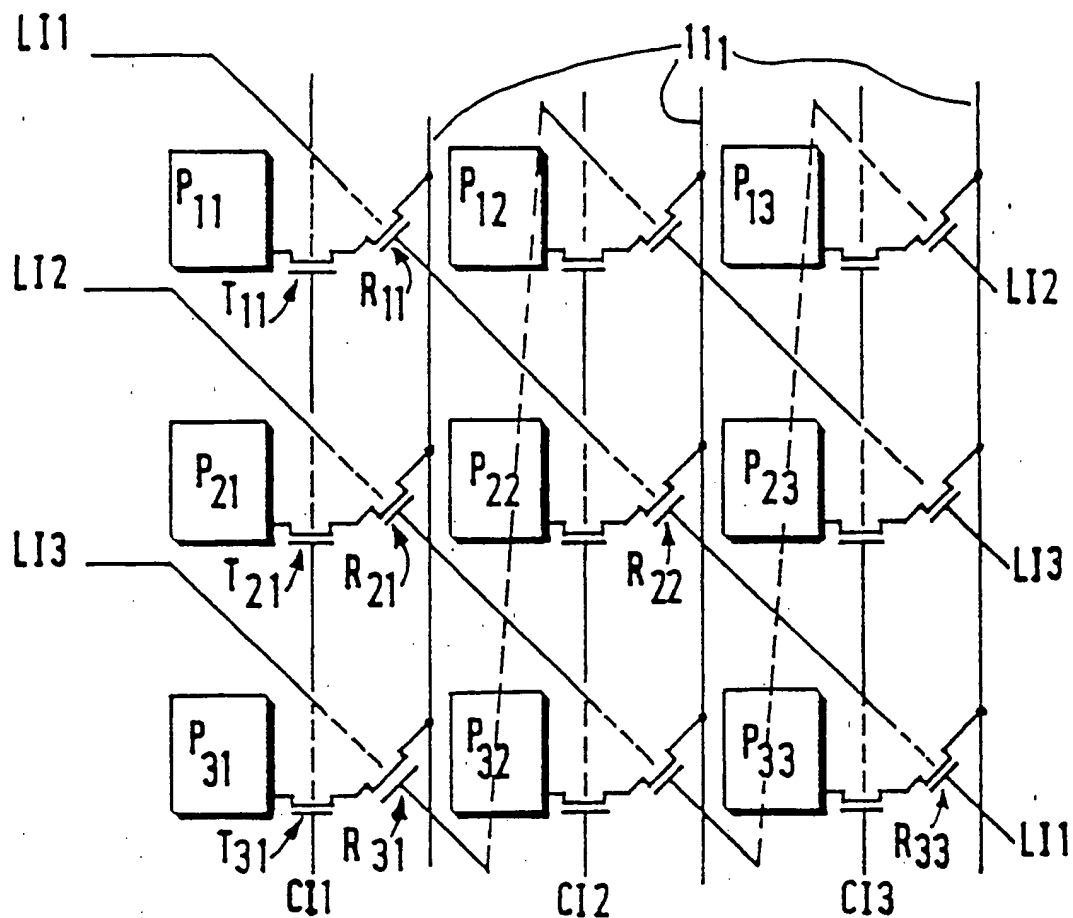


FIG. 4

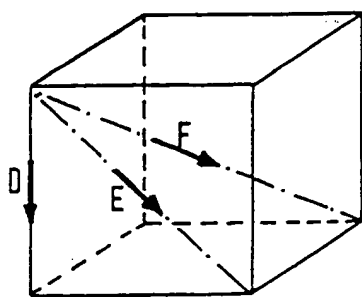


FIG. 6A

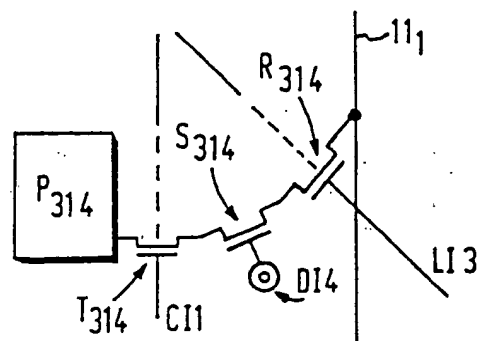
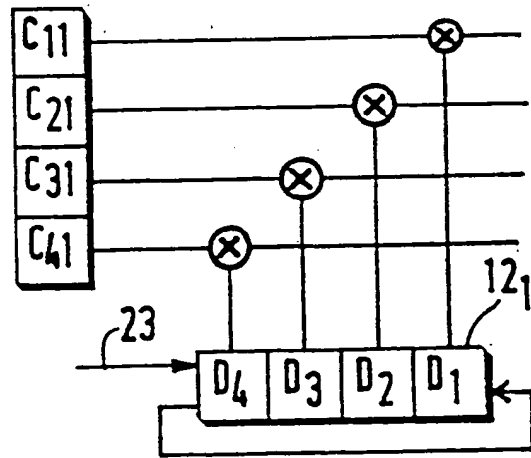
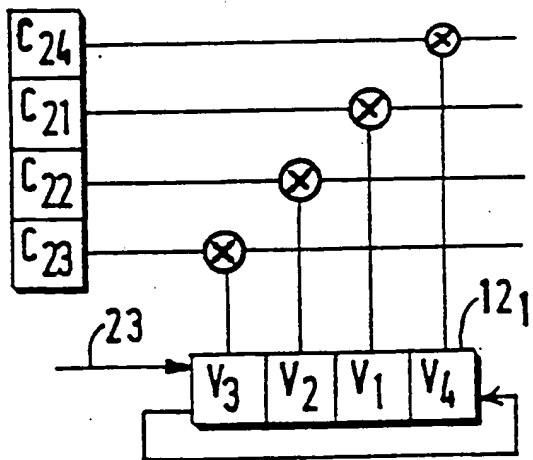
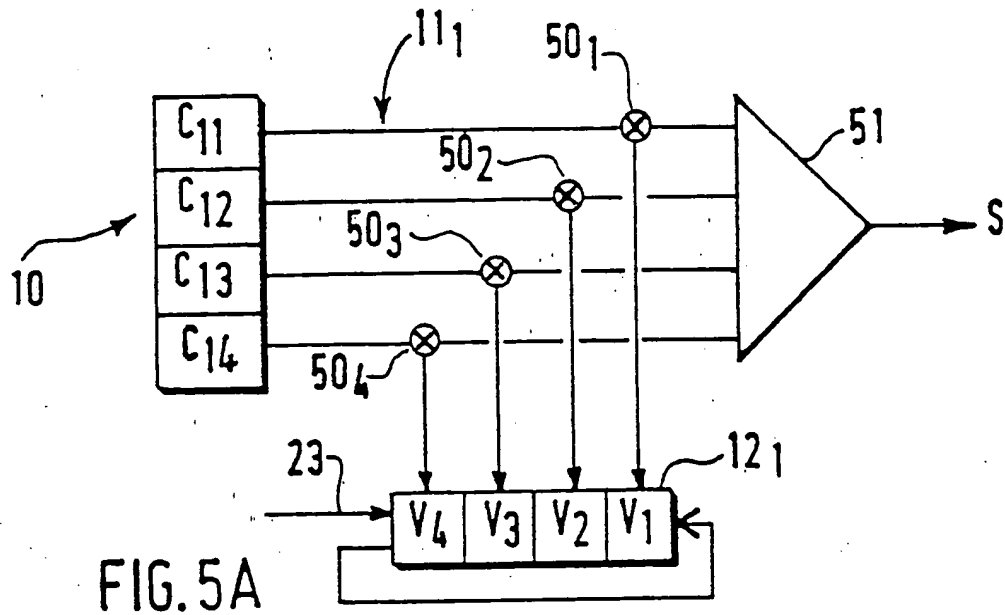


FIG. 6B





Office européen
des brevets

RAPPORT DE RECHERCHE EUROPEENNE

Numero de la demande

EP 89 20 2899

DOCUMENTS CONSIDERES COMME PERTINENTS			
Catégorie	Citation du document avec indication, en cas de besoin, des parties pertinentes	Revendication concernée	CLASSEMENT DE LA DEMANDE (Int. Cl.5)
A	US-A-4 287 571. (CHAKRAVARTI et al) * Résumé; figure 1; colonne 3, ligne 18 - colonne 2, ligne 25 *	1	G 11 C 8/00 G 11 C 5/02
A,D	WO-A-8 400 629 (MARCONI AVIONICS) * Résumé; page 2, lignes 15-23; figures 2-3; page 10, ligne 20 - page 14, ligne 12 *	1	
			DOMAINES TECHNIQUES RECHERCHES (Int. Cl.5)
			G 11 C 8/00 G 11 C 11/24 G 11 C 11/54 G 11 C 5/02
Le présent rapport a été établi pour toutes les revendications			
Lieu de la recherche LA HAYE		Date d'achèvement de la recherche 21-12-1989	Examinateur STECCHINA A.
CATEGORIE DES DOCUMENTS CITES			
X : particulièrement pertinent à lui seul Y : particulièrement pertinent en combinaison avec un autre document de la même catégorie A : arrière-plan technologique O : divulgation non-écrite P : document intercalaire		T : théorie ou principe à la base de l'invention E : document de brevet antérieur, mais publié à la date de dépôt ou après cette date D : cité dans la demande L : cité pour d'autres raisons & : membre de la même famille, document correspondant	

This Page Blank (uspto)